Supercomputer RAS via Informatics

(RAS – Reliability, Availability, Serviceability)

CIS External Review

August 19, 2004

Jon Stearley
jrstear@sandia.gov

Sandia National Laboratories
Dept. 9224

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company,
for the United States Department of Energy's National Nuclear Security Administration
under contract DE-AC04-94AL85000.

Plenty of Faults, Plenty of Logs, But Actionable Information is Elusive

- Event logs are a ubiquitous source of system feedback, but are notoriously free in format

- Supercomputers have many points of failure, with complex and dynamic interdependencies

∴ Identifying the root cause of faults in supercomputers is difficult (and expensive)

Leverage recent advances in informatics!

Impact

- Inspection of system logs is fundamental to debugging – increased capability to quickly extract meaningful information WILL impact MTTR (mean time to repair) and MAY impact MTBF (mean time between failure). Red Storm is principal impact target.

- An analysis system which accepts any time-stamped sequence of free-text messages will:
  - NOT be device specific: Computer, Network Switch; Linux, TOS, Catamount; Cplant, ASCI Red, Red Storm
  - NOT be application specific: RAS, security, others…

Charter Statement

With the specific goal of increasing supercomputer RAS, we intend to produce a machine-learning analysis system which enables content-novice analysts to efficiently understand evolving trends, identify anomalies, and investigate cause-effect hypotheses in large multiple-source log sets.

Natural Lanugage Analysts

Computer Security Analysts

Supercomputer Users

Systems Administrators

Automated Message Typing: Learning from Teiresias

What message content and occurrence rate is normal?

Teiresias (Bioinformatics code from IBM TJ Watson)

Two stages of operation:

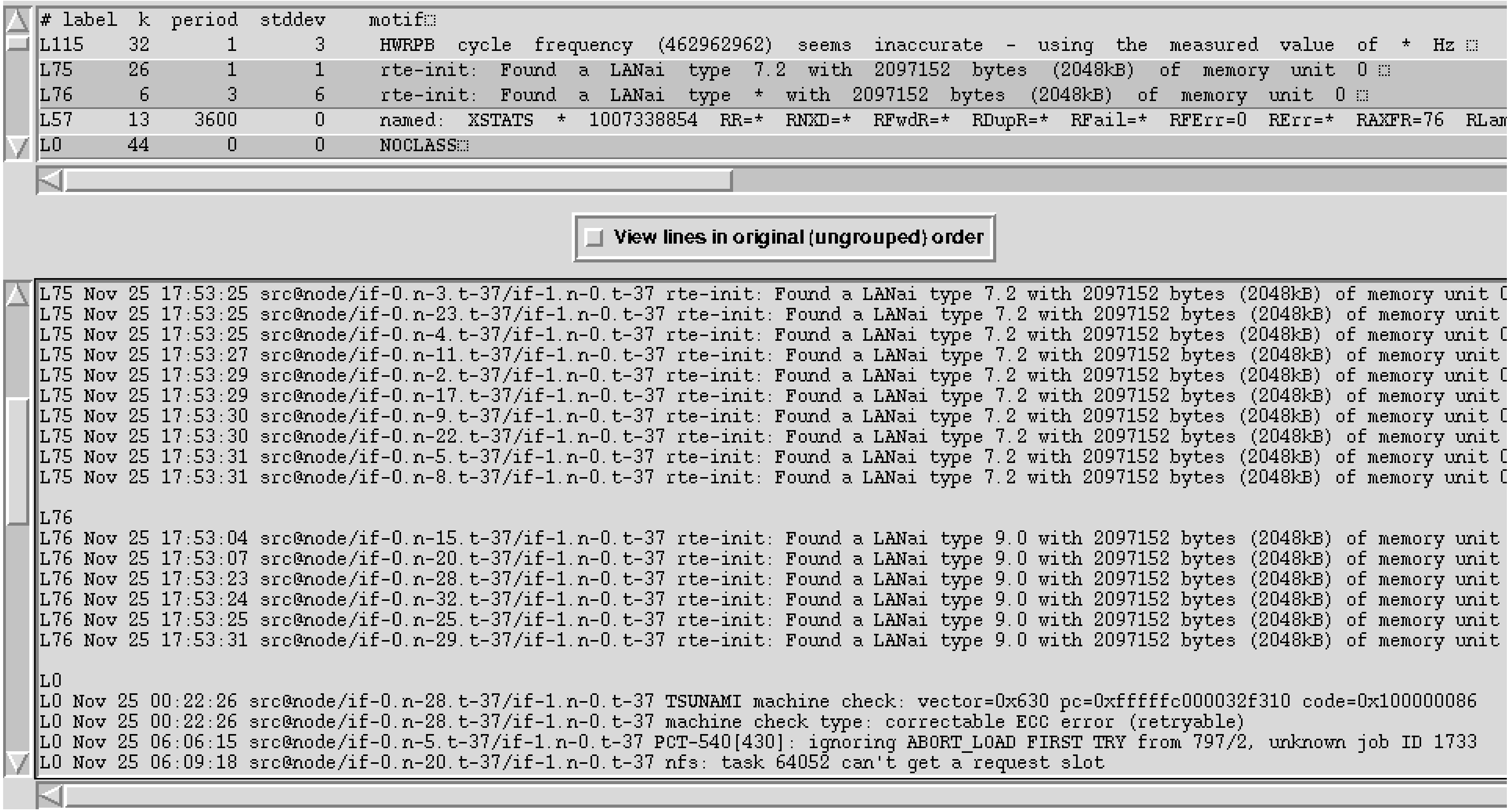
1. **Scanning**: enumerate all elementary patterns of at least L/W specificity (I.e. find all phrases of length W words, where at least L of them never change (the others do change))
2. **Convolution:** combine elementary patterns into maximal irredundant "motifs" (leverage property of "downward closure")

Sandia action: convert logs in/out of teiresias-acceptable format, cluster resulting "message templates" by time statistics

Result: automatically-generated "message templates", sorted into three content categories: common, deviant, anomalous

FY04 Progress: "Towards Informatic Analysis of Syslogs", IEEE Cluster '04

Interactive Review via Logview



```
# label  k  period  stddev    motif□
L115    32       1       3     HWRPB  cycle  frequency  (462962962)  seems  inaccurate  -  using  the  measured  value  of  *  Hz □
L75     26       1       1     rte-init:  Found  a  LANai  type  7.2  with  2097152  bytes  (2048kB)  of  memory  unit  0 □
L76      6       3       6     rte-init:  Found  a  LANai  type  *  with  2097152  bytes  (2048kB)  of  memory  unit  0 □
L57     13    3600       0     named:  XSTATS  *  1007338854  RR=*  RNXD=*  RFwdR=*  RDupR=*  RFail=*  RFErr=0  RErr=*  RAXFR=76  RLam
L0      44       0       0     NOCLASS□
```

□ View lines in original (ungrouped) order

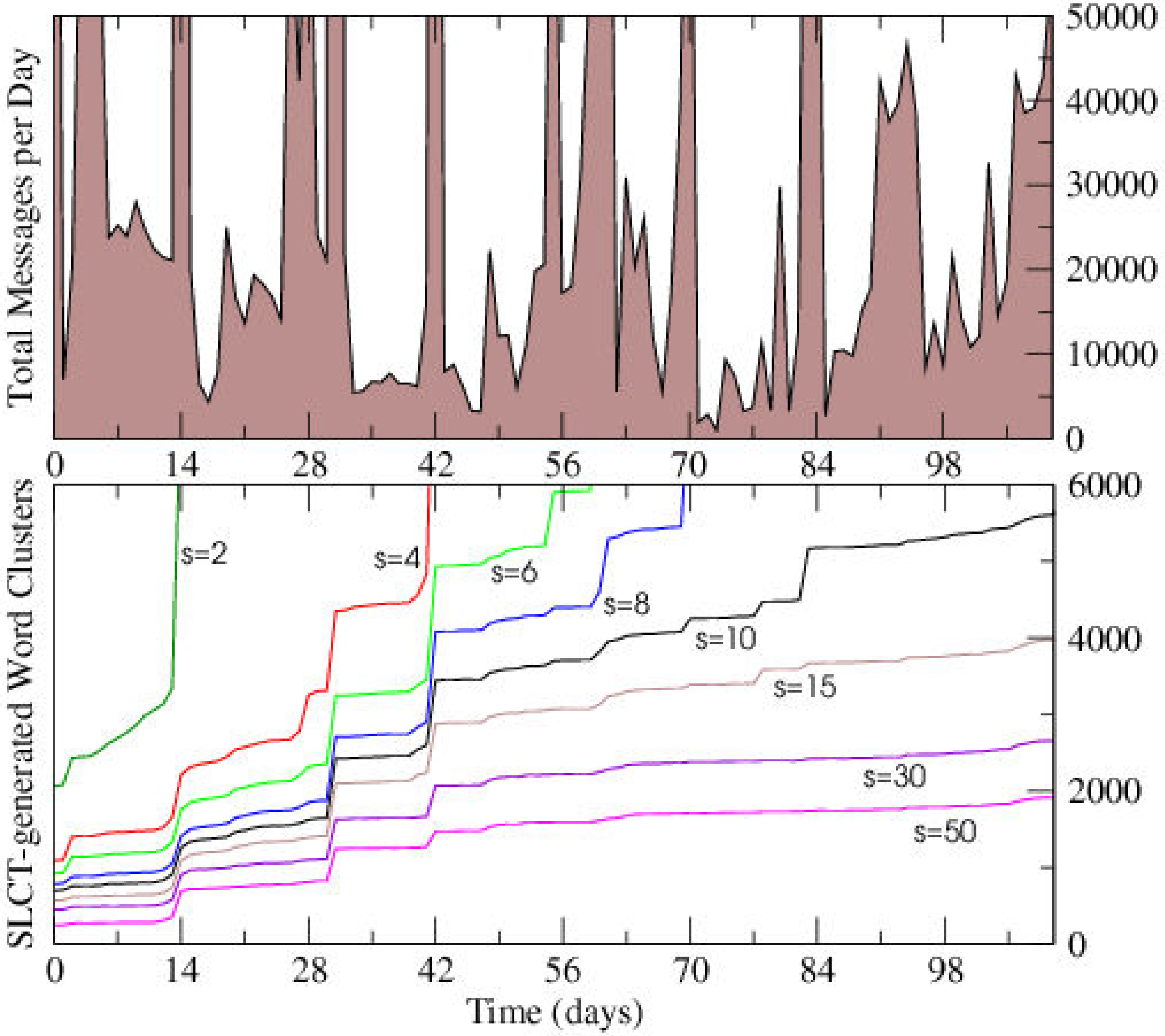
```
L75 Nov 25 17:53:25 src@node/if-0.n-3.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 7.2 with 2097152 bytes (2048kB) of memory unit 0
L75 Nov 25 17:53:25 src@node/if-0.n-23.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 7.2 with 2097152 bytes (2048kB) of memory unit
L75 Nov 25 17:53:25 src@node/if-0.n-4.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 7.2 with 2097152 bytes (2048kB) of memory unit 0
L75 Nov 25 17:53:27 src@node/if-0.n-11.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 7.2 with 2097152 bytes (2048kB) of memory unit
L75 Nov 25 17:53:29 src@node/if-0.n-2.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 7.2 with 2097152 bytes (2048kB) of memory unit 0
L75 Nov 25 17:53:29 src@node/if-0.n-17.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 7.2 with 2097152 bytes (2048kB) of memory unit
L75 Nov 25 17:53:30 src@node/if-0.n-9.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 7.2 with 2097152 bytes (2048kB) of memory unit 0
L75 Nov 25 17:53:30 src@node/if-0.n-22.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 7.2 with 2097152 bytes (2048kB) of memory unit
L75 Nov 25 17:53:31 src@node/if-0.n-5.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 7.2 with 2097152 bytes (2048kB) of memory unit 0
L75 Nov 25 17:53:31 src@node/if-0.n-8.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 7.2 with 2097152 bytes (2048kB) of memory unit 0

L76
L76 Nov 25 17:53:04 src@node/if-0.n-15.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 9.0 with 2097152 bytes (2048kB) of memory unit
L76 Nov 25 17:53:07 src@node/if-0.n-20.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 9.0 with 2097152 bytes (2048kB) of memory unit
L76 Nov 25 17:53:23 src@node/if-0.n-28.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 9.0 with 2097152 bytes (2048kB) of memory unit
L76 Nov 25 17:53:24 src@node/if-0.n-32.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 9.0 with 2097152 bytes (2048kB) of memory unit
L76 Nov 25 17:53:25 src@node/if-0.n-25.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 9.0 with 2097152 bytes (2048kB) of memory unit
L76 Nov 25 17:53:31 src@node/if-0.n-29.t-37/if-1.n-0.t-37 rte-init: Found a LANai type 9.0 with 2097152 bytes (2048kB) of memory unit

L0
L0 Nov 25 00:22:26 src@node/if-0.n-28.t-37/if-1.n-0.t-37 TSUNAMI machine check: vector=0x630 pc=0xfffffc000032f310 code=0x100000086
L0 Nov 25 00:22:26 src@node/if-0.n-28.t-37/if-1.n-0.t-37 machine check type: correctable ECC error (retryable)
L0 Nov 25 06:06:15 src@node/if-0.n-5.t-37/if-1.n-0.t-37 PCT-540[430]: ignoring ABORT_LOAD FIRST TRY from 797/2, unknown job ID 1733
L0 Nov 25 06:09:18 src@node/if-0.n-20.t-37/if-1.n-0.t-37 nfs: task 64052 can't get a request slot
```

FY04 Progress: "Towards Informatic Analysis of Syslogs", IEEE Cluster '04

Simple Logfile Clustering Tool (SLCT)

In contrast to Teiresias:

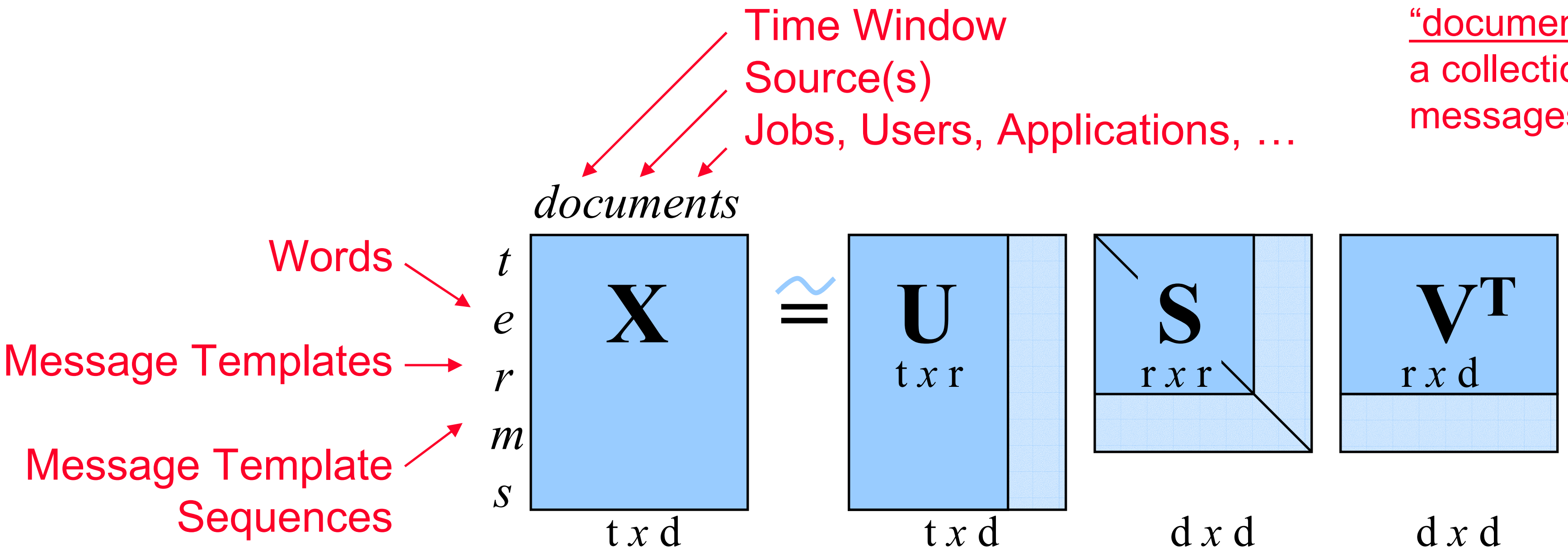
- Specifically designed to generate “message templates”
- Memory-efficient for log data
- Less effective anomaly categorization
- Open source ☺



FY04 Progress: "Towards Informatic Analysis of Syslogs", IEEE Cluster '04

Syslog Latent Semantic Analysis

Document similarity calculation using rank-reduced term space via Singular Value Decomposition.



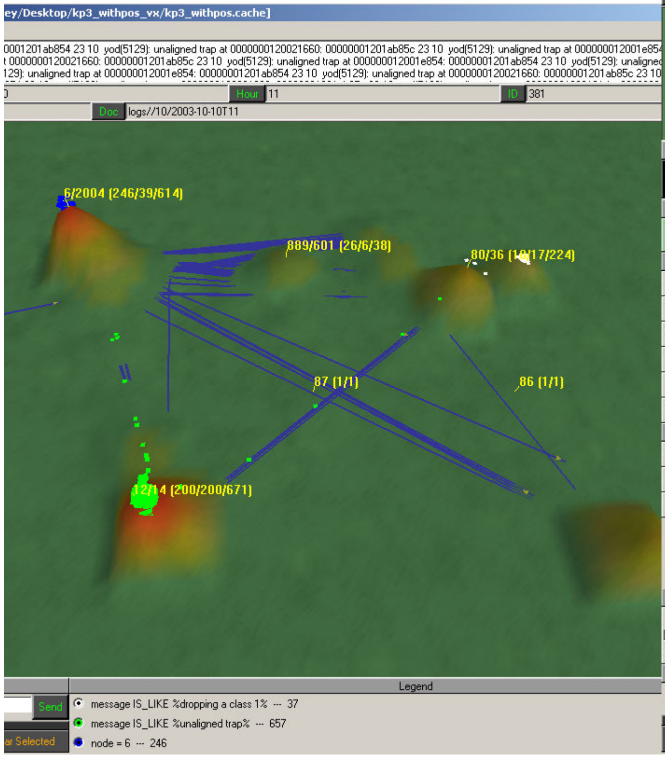
"document": a collection of messages

$$\text{Doc}\bullet\text{Doc similarity} = X^T X \cong X_r^T X_r \cong Y^T Y$$

$$\text{where } Y = \underset{\text{normalize}}{\text{column}}\left[S_r V_r^T \right]$$

Exploration via VxInsight

- Sandia (9212) application, uses include patent and gene research (and now, syslogs)
- SQL database underneath (provides flexible queries)



ey/Desktop/kp3_withpos_vx/kp3_withpos.cache]

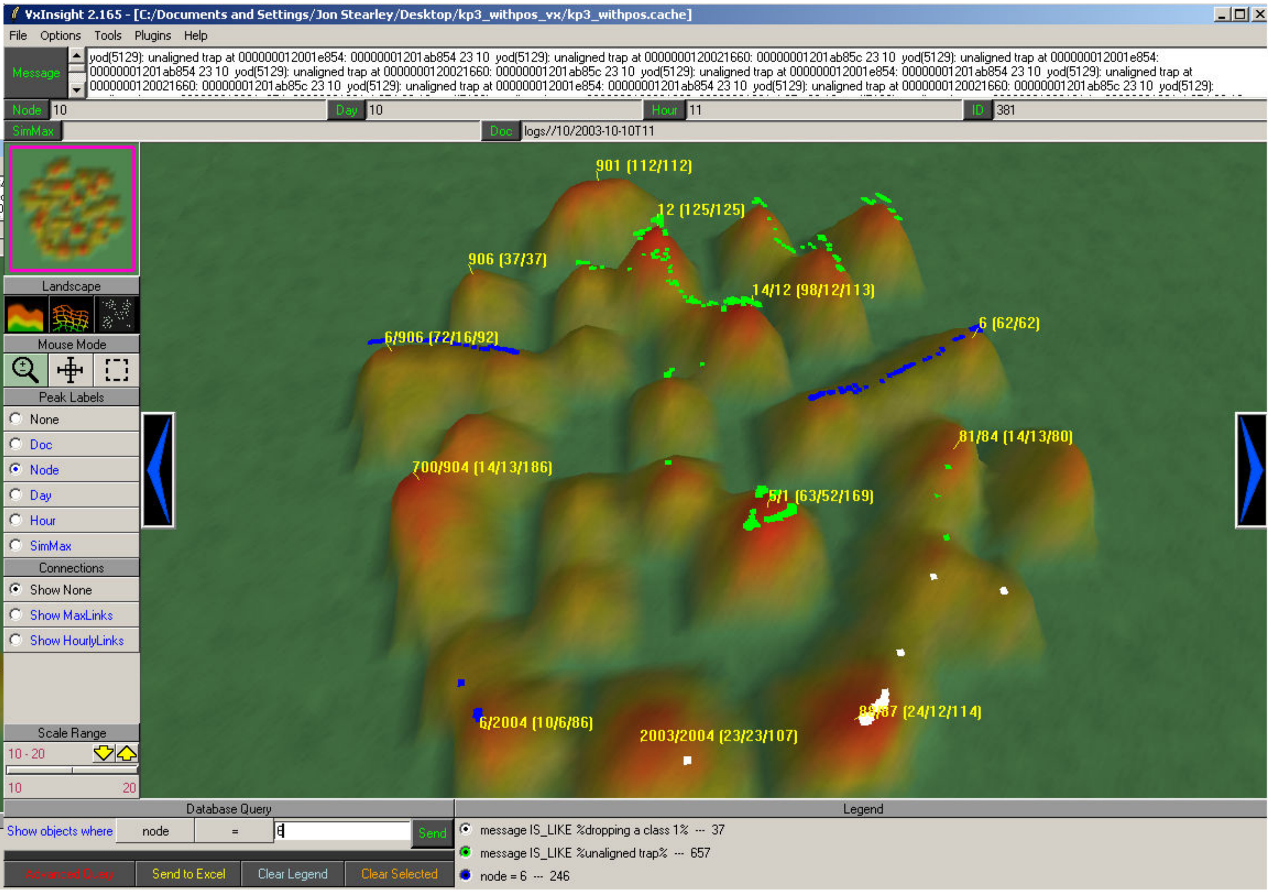
```
0001201ab854 23 10  yod(5129): unaligned trap at 000000012021660: 00000001201ab85c 23 10  yod(5129): unaligned trap at 00000001 2001e854
t 000000012021660: 00000001201ab85c 23 10  yod(5129): unaligned trap at 00000001 2001e854: 00000001201ab854 23 10  yod(5129): unaligned
129): unaligned trap at 00000001 2001e854: 00000001201ab854 23 10  yod(5129): unaligned trap at 000000012021660: 00000001201ab85c 23 10
```

Hour 11 ID 381

Doc logs//10/2003-10-10T11

Legend

- ○ message IS_LIKE %dropping a class 1% --- 37
- ● message IS_LIKE %unaligned trap% --- 657
- ● node = 6 --- 246



VxInsight 2.165 - [C:/Documents and Settings/Jon Stearley/Desktop/kp3_withpos_vx/kp3_withpos.cache]

Message:

```
yod(5129): unaligned trap at 00000001 2001e854: 00000001201ab854 23 10  yod(5129): unaligned trap at 000000012021660: 00000001201ab85c 23 10  yod(5129): unaligned trap at 00000001 2001e854:
00000001201ab854 23 10  yod(5129): unaligned trap at 000000012021660: 00000001201ab85c 23 10  yod(5129): unaligned trap at 00000001 2001e854: 00000001201ab854 23 10  yod(5129): unaligned trap at
000000012021660: 00000001201ab85c 23 10  yod(5129): unaligned trap at 00000001 2001e854: 00000001201ab854 23 10  yod(5129): unaligned trap at 000000012021660: 00000001201ab85c 23 10  yod(5129):
```

Node 10 | Day 10 | Hour 11 | ID 381

SimMax | Doc logs//10/2003-10-10T11

Peak Labels
- ○ None
- ○ Doc
- ● Node
- ○ Day
- ○ Hour
- ○ SimMax

Connections
- ● Show None
- ○ Show MaxLinks
- ○ Show HourlyLinks

Scale Range 10 - 20 (10 to 20)

Database Query: Show objects where node = 6 Send

Legend

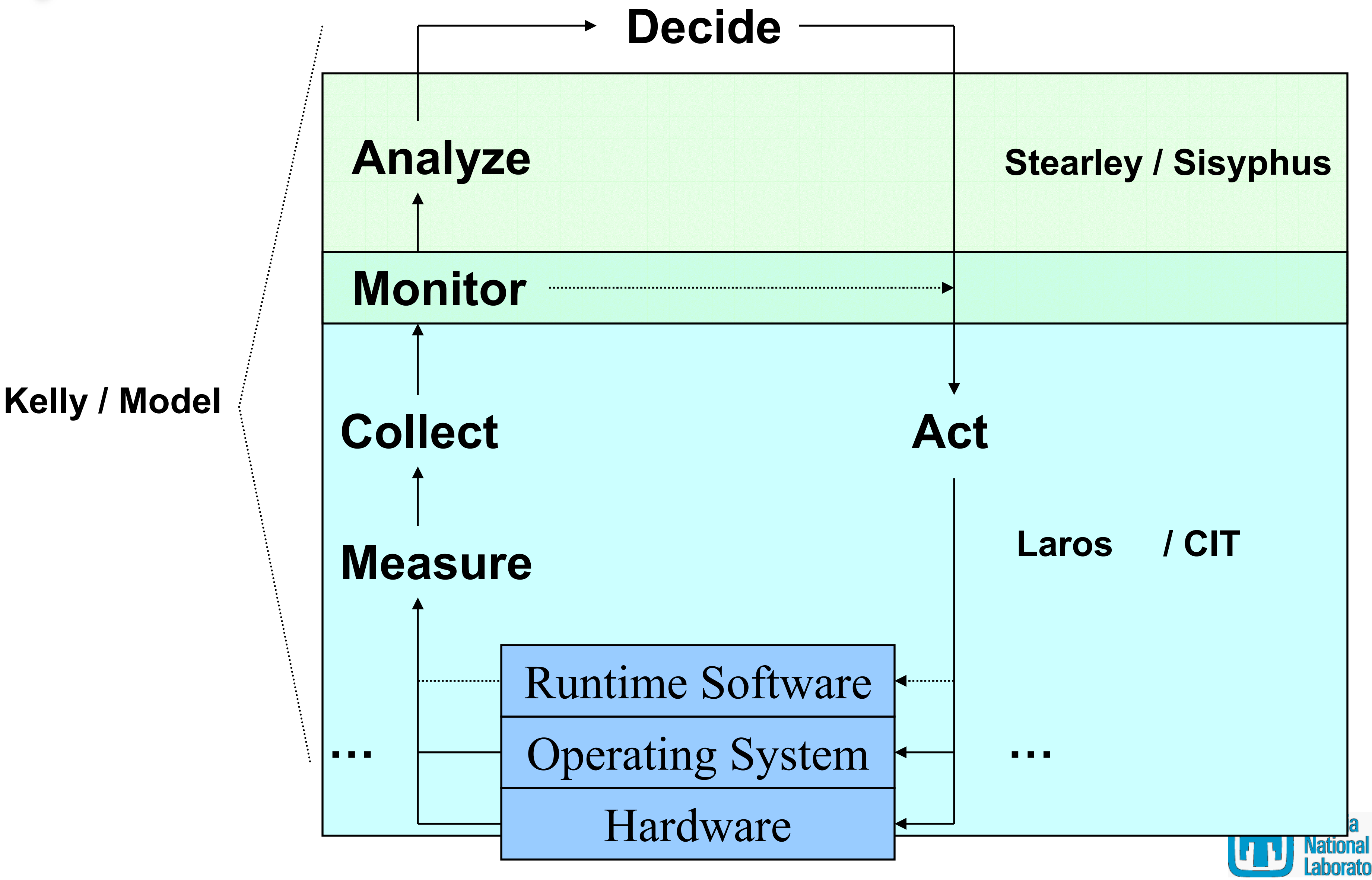
- ○ message IS_LIKE %dropping a class 1% --- 37
- ● message IS_LIKE %unaligned trap% --- 657
- ● node = 6 --- 246

Future Work

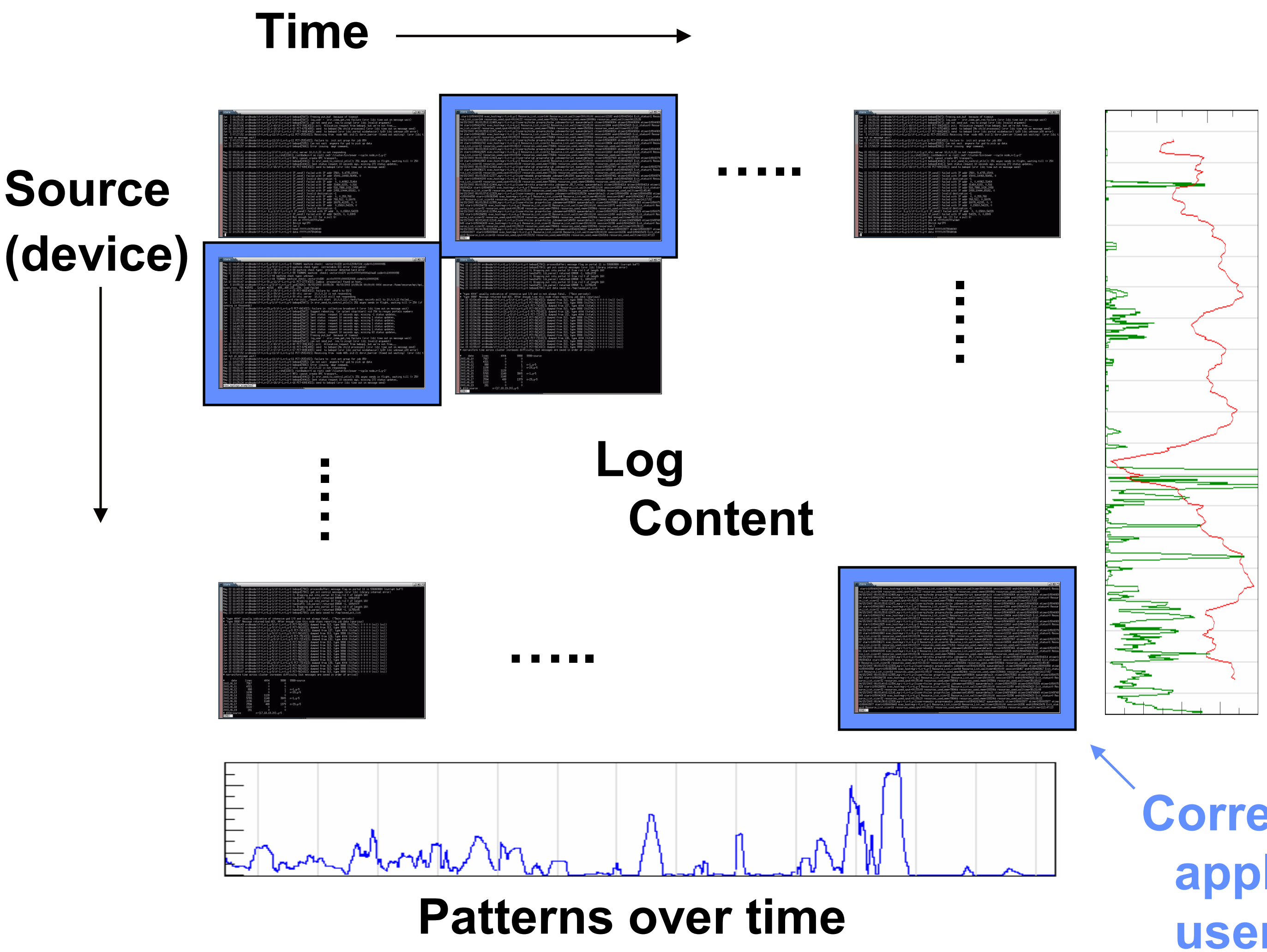
- Study novelty rate
  - Time rate of new message templates?
  - Rate of change of VxInsight landscape? (must reach "steady state" to be practical)
- Optimize parameters for effectiveness
  - "Term" and "Document" creation parameters?
  - Are traditional LSA weighting functions best for this application?
  - Degree of Rank Reduction?
- Quantitative effectiveness measure for test data sets
- Transition to RedStorm logs (RedStorm is primary impact target)
- Improve user interface
  - Expose "most distinguishing phrase" as VxInsight labels
  - Reduce manual effort required per parameter change; utilize Sandia Text Analysis Library "STANLEY" (Travis Bauer, 15241)

Backup Slides

RAS Projects Context



Supercomputer RAS Via Informatics



Time

Source (device)

Log Content

Patterns over time

Patterns in functional groups of devices, ie:

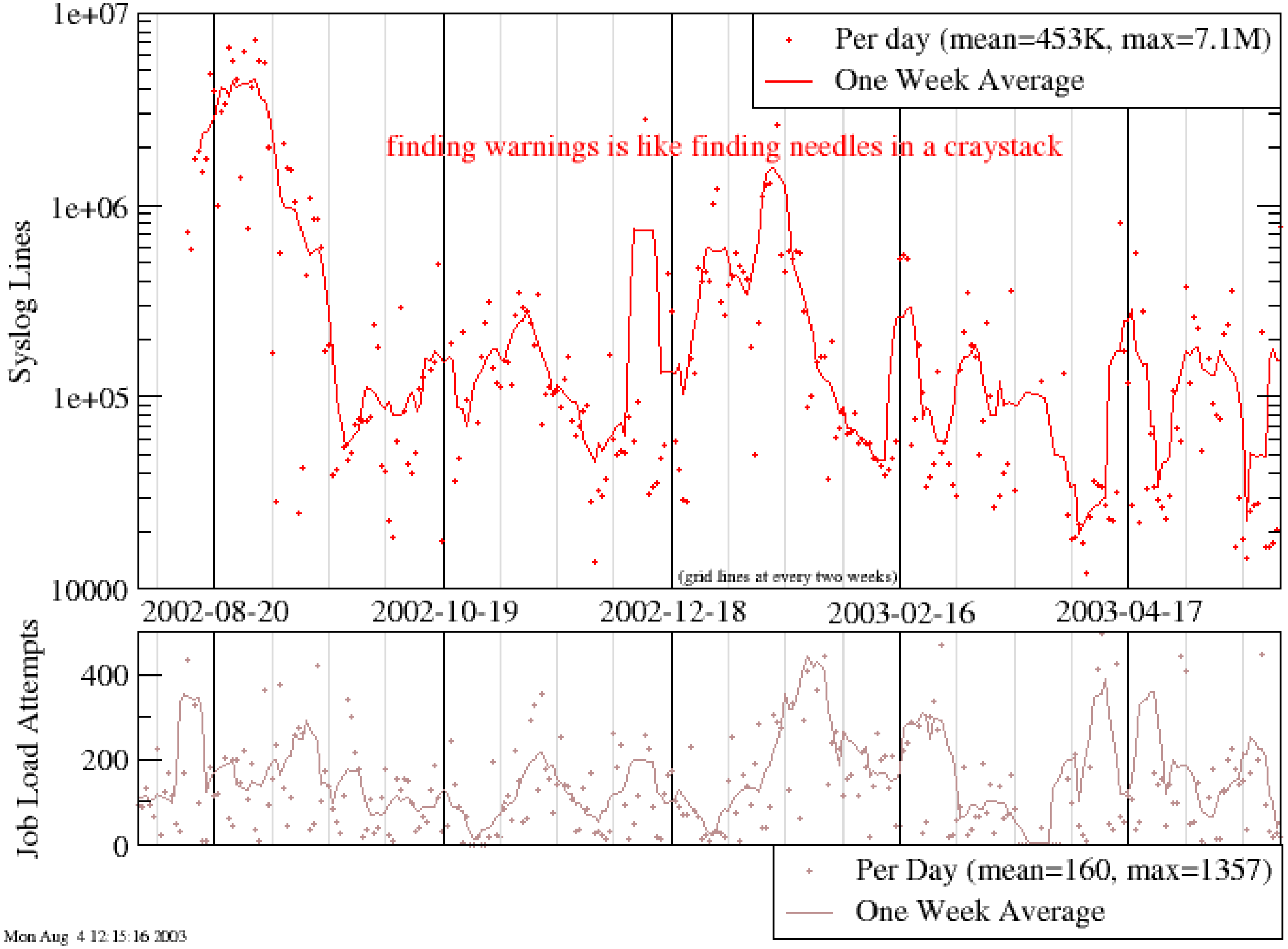
- Compute
- Login
- I/O
- Switches

Correlations with applications or users

Plentiful Data, Elusive Information

Ross Syslog Lines

(roughly periodic with biweekly system reboots, slight correllation with job turnover rate)



finding warnings is like finding needles in a craystack

Mon Aug 4 12:15:16 2003

Research Context

- Computer Security - misuse and intrusion detection
- Information Mgmt - search engines, translation
- Health - gene and protein sequencing
- National Security - language modeling, antiterrorism

Nobody is leveraging informatics towards supercomputer RAS

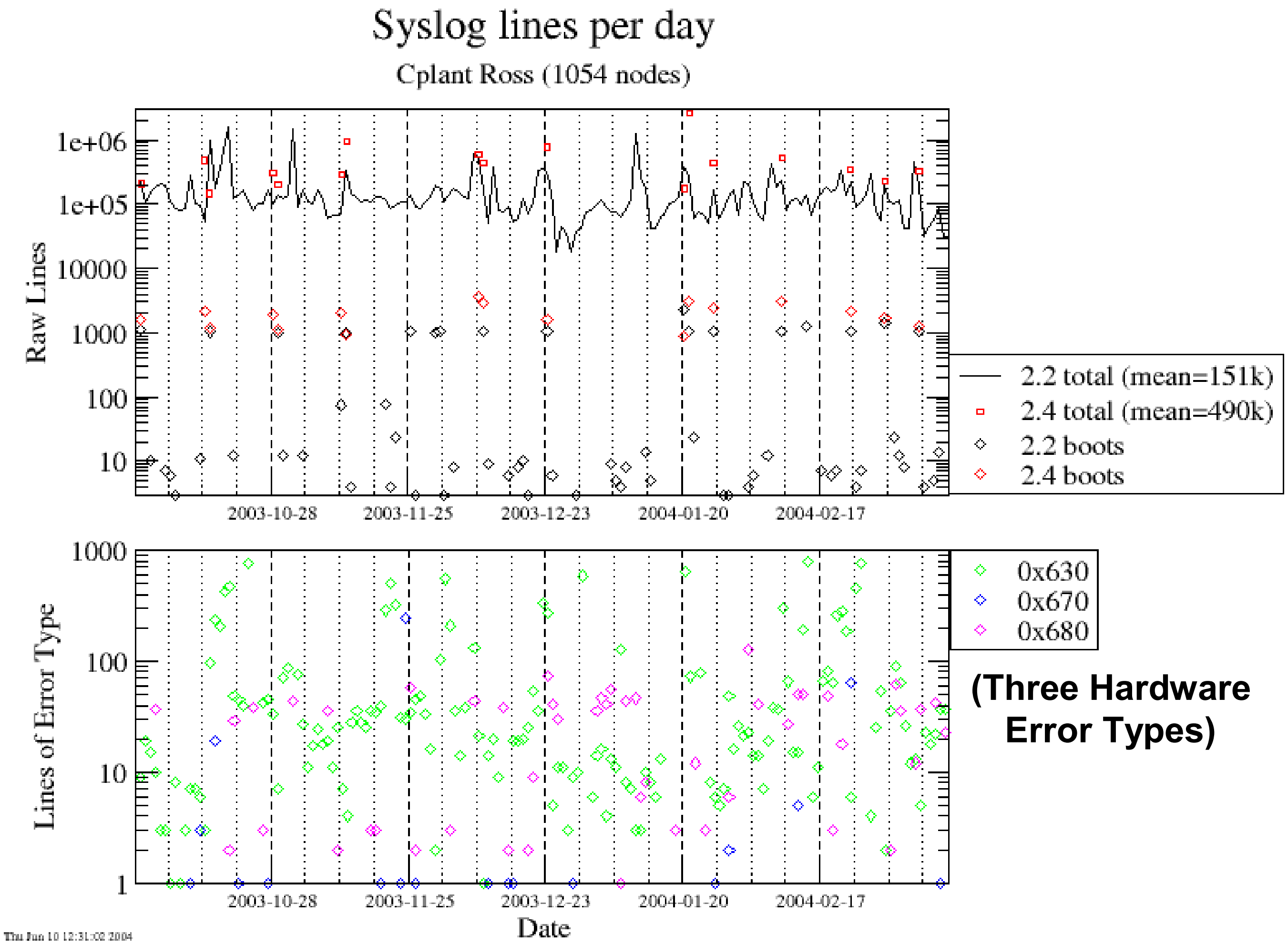
(upon which many of the above depend)

Analyst Thought Process

"What aspects of the message stream are strongly correlated with system malfunction or misuse?"

1. What message content and occurrence rate is normal?
2. What message groups are normal?
3. Can devices be classified by their output message stream?
4. Can users or applications be classified by their resulting message stream?
5. Are device-to-device and/or job-to-job log stream similarities sufficient to identify hardware or software failures?

Fault-Annotated Log Database



(Three Hardware Error Types)

SQL Database supports flexible subset selection:

- By Device
- By Time
- By Job
- By User
- By Error or Message Type
- etc

Automated Grouping of Time-correlated Messages

What message groups are normal?

Approach:

Cluster message templates using their occurrence statistics:

1. Support
2. Inter-arrival median
3. Inter-arrival standard deviation

Results:

1. Detects periodic messages
2. Groups message groups which are rigidly time-correlated.

Automated Message Typing: Leveraging SLCT

SLCT – Simple Logfile Clustering Tool

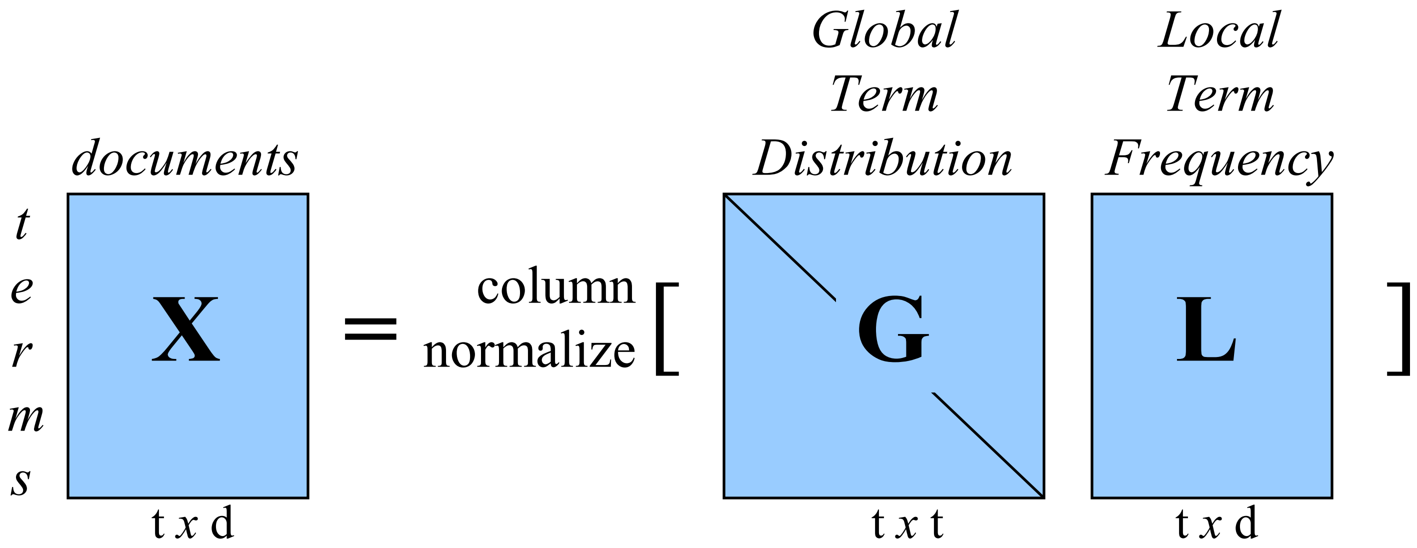
Operates in three phases:

1. Frequent words –count all {position, word} tuples (pw), prune those occurring less than S times
2. Frequent messages – count all single-line pw "clusters" ("1w 2w 3_* 4_w", ie message types)
3. Wildcard refinement (optional) – determine constant prefix or suffix for wildcards

Output and Categorization:

- pw clusters occurring at least S times (**common**)
- less specific pw clusters which, if joined with above pw clusters, occur at least S times (**common+deviant**)
- lines not matching either of the above (**rare**)

Term-Document Matrix



documents

terms

X

t *x* d

=

column normalize [

Global Term Distribution

G

t *x* t

Local Term Frequency

L

t *x* d

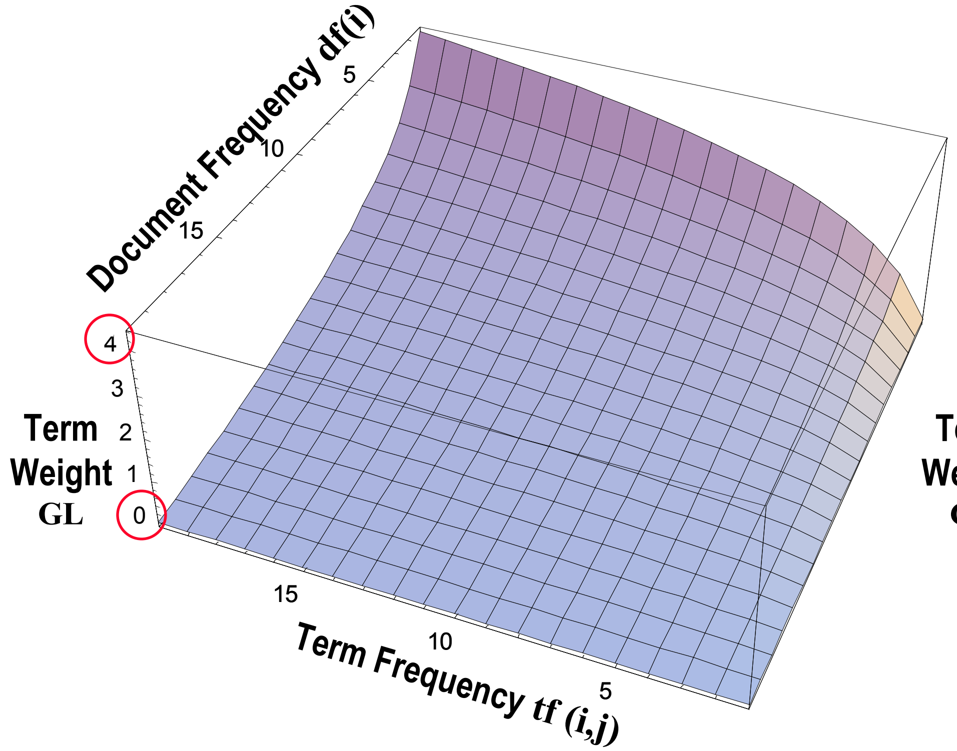
]

G and **L** are weighting functions.

Most simple case is **G=I** and **L=tf(i,j)**, where **tf(i,j)** is “term-frequency” of **i**’th term in **j**’th document)

Term-Doc Matrix Weighting Functions

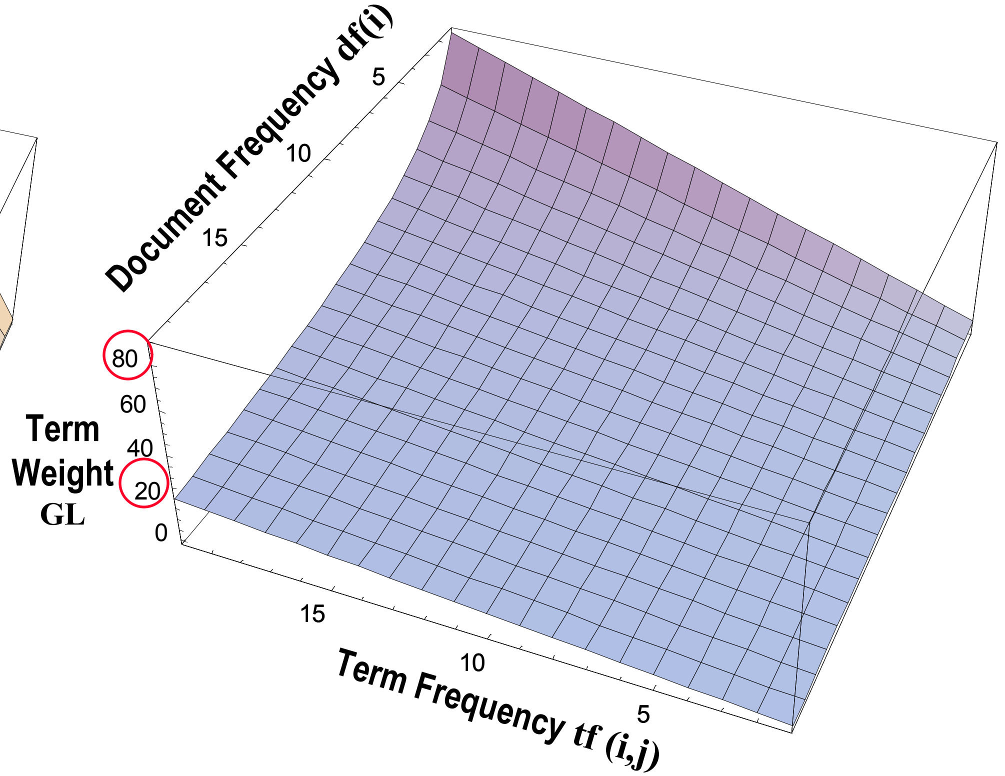
Log-Entropy



$$L(i,j) = \log\left(tf(i,j)+1\right)$$

$$G(i) = 1 + \frac{1}{\log(n)} \sum_{j=1}^{n} p_{ij} \log(p_{ij}), \quad p_{ij} = \frac{tf(i,j)}{\sum_{j=1}^{n} tf(i,j)}$$

Inverse Document Frequency



$$L(i,j) = tf(i,j)$$

$$G(i) = \log\left(\frac{n}{df(i)} + 1\right)$$